

State Price Densities implied from weather derivatives

Wolfgang Karl Härdle*
Brenda López-Cabrera*
Huei-Wen Teng**



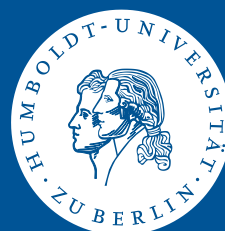
* Humboldt-Universität zu Berlin, Germany

** Graduate Institute of Statistics National Central University,
Taiwan

This research was supported by the Deutsche
Forschungsgemeinschaft through the SFB 649 "Economic Risk".

<http://sfb649.wiwi.hu-berlin.de>
ISSN 1860-5664

SFB 649, Humboldt-Universität zu Berlin
Spandauer Straße 1, D-10178 Berlin



State Price Densities implied from weather derivatives

*

Wolfgang Karl Härdle[†], Brenda López-Cabrera[‡], Huei-Wen Teng[‡]

March 18, 2013

Abstract

A State Price Density (SPD) is the density function of a risk neutral equivalent martingale measure for option pricing, and is indispensable for exotic option pricing and portfolio risk management. Many approaches have been proposed in the last two decades to calibrate a SPD using financial options from the bond and equity markets. Among these, non and semi parametric methods were preferred because they can avoid model mis-specification of the underlying and thus give insight into complex portfolio propelling. However, these methods usually require a large data set to achieve desired convergence properties. Despite recent innovations in financial and insurance markets, many markets remain incomplete and there exists an illiquidity issue. One faces the problem in estimation by e.g. kernel techniques that there are not enough observations locally available. For this situation, we employ a Bayesian quadrature method because it allows us to incorporate prior assumptions on the model parameters and hence avoids problems with data sparsity. It is able to compute the SPD of both call and put options simultaneously, and is particularly robust when the market faces the illiquidity issue. By comparing our approach with other approaches, we show that the traditional way of estimating the SPD by differentiating an interpolation of option prices does not hold in practice. As illustration, we calibrate the SPD for weather derivatives, a classical example of incomplete markets with financial contracts payoffs linked to non-tradable assets, namely, weather indices. Finally, we study the dynamics of the implied SPD's and related to weather data.

Keywords: Weather derivatives, temperature derivatives, HDD, CDD, SPD, mixture, quadrature, Bayesian, Option trading Strategies, illiquid

JEL classification: C11, C22, C58, G12, G13, G19, G22, N23, N53

*The financial support from the Deutsche Forschungsgemeinschaft via SFB 649 “Ökonomisches Risiko”, Humboldt-Universität zu Berlin is gratefully acknowledged. The third author was supported in part by the National Science Council of Taiwan under Grant NSC 101-3113-P-008-005

[†]Humboldt-Universität zu Berlin, Ladislaus von Bortkiewicz Chair of Statistics, Spandauer Straße 1, 10178 Berlin, Germany, Emails: haerdle@wiwi.hu-berlin.de, lopezcab@wiwi.hu-berlin.de

[‡]Graduate Institute of Statistics National Central University, No. 300, Jhongda Rd., Jhongli City, Taoyuan County 32001, Taiwan (R.O.C.). Email: wenteng@ncu.edu.tw

1 Introduction

A State Price Density (SPD) is the density function of a Risk Neutral (RN) equivalent martingale measure for option pricing, it is a measure more tied to uncertainty than to volatility and it is indispensable for (exotic) option pricing and portfolio risk management. It does not only reflect a risk-adaptive behaviour of investors based on historical assessment of the futures market, but it also gives insights about the preferences and risk aversion of a representative agent, see for example Ait-Sahalia and Lo (2000), Jackwerth and Rubinstein (1996), Rosenberg and Engle (2002).

Consider a European call option with maturity date T and strike price K . Under the non-arbitrage principle, its price at t can be given as:

$$C(K) = e^{-r\tau} \int \max(x - K, 0) f(x) dx \quad (1)$$

where r is the risk-free interest rate, τ time to maturity and $f(x)$ is the defined SPD. The advantages of extracting the SPD directly from market prices is that volatility and other moments can easily be calculated using this SPD independent of any particular pricing model.

There are any approaches to calibrate the SPD using financial options from the bond and equity markets. Assuming a Black and Scholes (B&S) model implies that the RN measure is a lognormal distribution which may result in SPD estimation since certain volatility properties are not correctly reflected. As observed by Breeden and Litzenberger (1978), the SPD of any risky asset can be derived as the second derivative with respect to the strike price of an estimate of the pricing function C . A number of econometric techniques have been developed to address this calibration issue. The most notable examples include the stochastic volatility models and the GARCH models. Derman and Kani (1994), Dupire (1994), Rubinstein (1994) implied SPDs using binomial trees, hence avoiding too strong stochasticity assumption like e.g. Geometric Brownian motion. Others like Abadir and Rockinger (2003) use hypergeometric distributions. Although useful in a variety of contexts, these (parametric) models are still susceptible to model specification.

Various non-parametric models have been employed to overcome this problem. Ait-Sahalia and Lo (1998) introduce a semiparametric alternative where the volatility of the B&S formulation is modelled non-parametrically. From a statistical point of view, estimating the SPD becomes estimating the second derivative of a regression function, but the SPD needs to be a proper density function (non negative and integrates to one). This dictates that the price is monotone decreasing and convex in terms of the strike price. How to impose these constraints presents the main difficulties of direct applications of nonparametric regression. Ait-Sahalia and Duarte (2003), Yatchew and Härdle (2006) and Härdle and Hlávka (2009) stress the importance of enforcing such shape constraints. Fan and Mancini (2009) use a non-parametric technique to estimate the state price distribution but not the density because the former is easier to estimate. Giacomini et al. (2008) use mixtures of scales and shifted t -distributions, while Yuan (2009) uses a mixture of lognormals. Curve fitting method have been presented in Rubinstein (1994), Jackwerth and Rubinstein (1996). Teng and Liechty (2009) introduce the Bayesian quadrature model, which is a generalization of the method of Rubinstein (1994) in the sense that both

the location and weights of the support points are random variables. Most nonparametric methods require a rich body of data to achieve desired convergence properties. The main goal of this paper is to infer the SPD from markets, where there is little liquidity of the assets.

For this purpose, we employ a Bayesian quadrature method as a calibration method for the SPD from option prices, because it allows us to incorporate prior assumptions on the model parameters and hence avoids problems with data sparsity. This approach takes a prior distribution which can be parametric (e.g. lognormal) or a uniform density. The posterior distribution of the SPD is calibrated to market data. This method is a special case of a mixture model, where the component densities are point measures.

We show that the proposed method has some advantages over other nonparametric methods: 1) it considers the locations and weights of the support points in the finite representation of the SPD as random variables, 2) it is parsimonious and allows for statistical inference, it enables us to construct credible regions for the current value of the SPD 3) it is computationally efficient in the sense that a Markov chain Monte Carlo algorithm with Gibbs sampler can be adopted, so that no additional tuning procedures are required for exploring the posterior distribution and 4) it is robust even if the market faces illiquidity issues. 5) We show that the traditional way to estimate the SPD by differentiating an interpolation of option prices does not hold in practice.

We conduct our empirical analysis based on weather derivative (WD) data traded at the Chicago Mercantile Exchange (CME). WDs are newly developed financial instruments. Key features of weather derivatives are that the underlying process, i.e., temperature index, is not tradeable. Consequently, the Black-Scholes formula is unsuitable since an essential element of it is the tradability of the underlying. In addition, the temperature index shows apparent seasonality and it is determined by physical phenomena. An interesting feature is that weather futures and options are rarely traded and traded only at a few strike prices compared with other liquid equity markets. The CME (the official WD platform) provides closing prices, which are however not the real trading prices negotiated by the market participants. The SPD enables to price options with complicated payoff functions simply by numerical integration of the payoff with respect to this density. However, data sparsity makes the SPD estimation a statistical challenge.

The estimated curves are compared with other popular SPD density estimation methods (lognormal, kernel density, mixture of log-normal). In addition we study the dynamics of the SPD which provides useful insight into the economic behavior of agents sensitive to weather conditions and the time inhomogeneity of the market.

This paper is structured as follows. Section 2 describes the quadrature approach and its comparison to other popular SPD density estimation methods. In Section 3, we address the illiquidity issue by addressing why other nonparametric methods fail particularly when options with only a few strike prices are traded. Section 4 conducts the empirical analysis of SPDs from CME weather option data using different methods and study the dynamics of the SPD weather type. Section 5 concludes the paper. All quotations of currency in this paper will be in USD and therefore we will omit the explicit notion of the currency. All the SPDs computations were carried out in Matlab version 7.6. The option data on temperature indices were obtained from CME and are also available from the research

data center of the CRC 649 "Economic Risk".

2 The Bayesian quadrature method

Options are contingent claims on an underlying asset. Plain vanilla option are of either put or call type with a fixed maturity, i.e., the value of the underlying is compared to a strike price K at maturity T . Let x denote the underlying asset's price at maturity (in our application this will be equivalent to futures prices on weather indexes). For a call option, one has the payoff $\max(x - K, 0)$ and for a put $\max(K - x, 0)$. If we denote a put as $i = 1$ and a call with $i = 2$, and observed strike prices K_{ij} for $i = 1, 2$ and $j = 1, \dots, N_i$ indexing all possible strike prices on any given day t , then the payoff function at maturity, denoted by $\wp_{ij}(x)$, can be represented by one formula,

$$\wp_{ij}(x) = (-1)^i(x - K_{ij})\mathbf{I}\{(-1)^i(x - K_{ij}) > 0\}(x),$$

where $\mathbf{I}\{A\}$ is an indicator function for a set A . Let t be the current time. The fair option price is given as (1) as the discounted value of the expected payoff function:

$$C_{ij} = \exp(-r\tau) \mathbf{E}^Q[\wp_{ij}(x)],$$

where $\tau = T - t$ is the time to maturity and $\mathbf{E}^Q[\cdot]$ is the expectation operator taken under the risk-neutral measure. The density $f(x)$ under this risk-neutral measure is the defined SPD. When the SPD $f(x)$ exists, this equals:

$$C_{ij} = \exp(-r\tau) \int \wp_{ij}(x) f(x) dx. \quad (2)$$

The left hand side of (2) is observed on the market for different payoff types depending on put/call ($i = 1, 2$), strike price K_{ij} , and time to maturity τ . The interest of statistical calibration is to infer the SPD $f(x)$ from a set of observed option prices.

2.1 The quadrature method

The word "quadrature" means a numerical method to approximate an integral either analytically or numerically, see Ueberhuber (1997) for example. In this research, we work the adverse way, since the interest is to infer the unknown density from the observed integrals (options prices). Define the δ -function $\delta_{\varpi}(\cdot)$ as a unit point measure at the location s by

$$\delta_s(x) = \mathbf{I}\{s = x\}$$

The basic idea of the quadrature method is to approximate the SPD $f(x)$ by $f_N(x|w, \theta)$, a weighted sum of δ -functions:

$$f_N(x|w, \theta) = w_1\delta_{\theta_1}(x) + \dots + w_N\delta_{\theta_N}(x), \quad (3)$$

with unknown locations $\theta = (\theta_1, \dots, \theta_N)^\top$ and weights $w = (w_1, \dots, w_N)^\top$. Here, N is a non-negative integer (smoothing) parameter. To produce a legitimate probability

density, the locations θ are constrained to be non-negative quantities, and the weights w are constrained to be nonnegative quantities and sum up to one. From a modeling perspective, the quadrature method (3) can be seen as a finite mixture distribution with the point measure as the component density. Figure 1 illustrates (3) for $N = 5$.

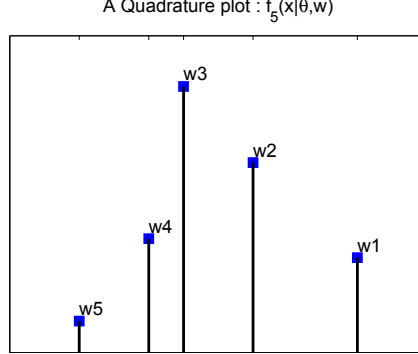


Figure 1: The SPD $f_N(x|w, \theta)$ from (3) for $N = 5$

The option price (2) under $f_N(x|w, \theta)$ is:

$$C_{ij}^N(w, \theta) = \exp(-r\tau) \sum_{n=1}^N w_n \phi_{ij}(\theta_n). \quad (4)$$

Note that (4) is an approximation to (2) and the aim of calibration is to extract (w, θ) by matching $C_{ij}^N(w, \theta)$ to the observed option prices. More specifically, a call option price calculated with (3) is:

$$C_{2j}^N(w, \theta) = \exp\{-r\tau\} \sum_{n=1}^N w_n \max(\theta_n - K_{ij}, 0), \quad (5)$$

whereas a put option price under the quadrature method is:

$$C_{1j}^N(w, \theta) = \exp\{-r\tau\} \sum_{n=1}^N w_n \max(K_{ij} - \theta_n, 0). \quad (6)$$

2.2 Bayesian modeling and computation

Empirical observations show that options having higher prices usually have higher price variation, see Ghysels et al. (1995) and Ghysels et al. (1997). Hence for the calibration task as a variance stabilizing transformation, we consider the logarithm of option prices. The observations y_{ijk} are perturbations of the model option price $C_{ij}^N(w, \theta)$:

$$\log y_{ijk} = \log C_{ij}^N(w, \theta) + \varepsilon_{ijk} \quad (7)$$

for $i = 1, 2, j = 1, \dots, N_i, k = 1, \dots, N_{ij}$, where the error $\varepsilon_{ijk} \sim \mathcal{N}(0, \sigma^2)$. ε_{ijk} is attributed to market friction and the approximation discrepancy, Garcia et al. (2010) and Renault

(1997). In Section 3, residual analysis of our empirical studies will support this error assumption. Following (7), the likelihood is

$$L(y|w, \theta, \sigma^2) = \prod_{i=1}^2 \prod_{j=1}^{N_i} \prod_{k=1}^{N_{ij}} (2\pi\sigma^2)^{-\frac{1}{2}} \exp \left[-\frac{\{\log y_{ijk} - \log C_{ij}^N(w, \theta)\}^2}{2\sigma^2} \right] \quad (8)$$

A natural prior distribution for the weights w is the Dirichlet distribution, which ensures w being positive and summing up to one. The Dirichlet distribution with parameter $\gamma = (\gamma_1, \dots, \gamma_N)^\top$ has the density function,

$$f(w|\gamma) = \frac{1}{B(\gamma)} \prod_{n=1}^N w_n^{\gamma_n-1} \quad (9)$$

for $w_n > 0$, $n = 1, \dots, N$, and $w_1 + \dots + w_N = 1$. The normalizing constant $B(\gamma)$ is defined as

$$B(\gamma) = \frac{\prod_{n=1}^N \Gamma(\gamma_n)}{\Gamma(\sum_{n=1}^N \gamma_n)}$$

where $\Gamma(\cdot)$ is the gamma function, Chen and Shao (1997).

Let K_{\min} and K_{\max} denote the minimum and maximum of the observed strike prices K_{ij} , respectively. To avoid label switching problems for θ , we assume that the locations are ordered, i.e., $\theta_1 \leq \dots \leq \theta_N$. Moreover, to avoid model option prices in (4) being zeros, assume a priori that the smallest location, θ_1 , is less than the minimum of the observed strike price, and that the largest location, θ_N , is larger than the maximum of the observed strike prices. Therefore, we assume that the distribution of the locations θ are uniformly distributed over the set $\{\theta_1 \leq \theta_2 \leq \dots \leq \theta_N, \theta_1 < K_{\min}, \theta_N > K_{\max}\}$:

$$f(\theta|K_{\min}, K_{\max}) \propto \mathbf{I}\{\theta_1 \leq \dots \leq \theta_N, \theta_1 < K_{\min}, \theta_N > K_{\max}\}(\theta) \quad (10)$$

For simplicity, we consider an inverse-gamma distribution with shape parameter α and scale parameter β as a prior distribution for σ^2 , denoted by $\sigma^2 \sim IG(\alpha, \beta)$. The prior density of σ^2 is

$$f(\sigma^2|\alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} (\sigma^2)^{-\alpha-1} \exp \left(-\frac{\beta}{\sigma^2} \right) \quad (11)$$

Putting things together allows a conjugate prior for σ^2 , as described in Casella and Berger (2001).

Note that (9-11) can be changed in cases where appropriate information is available. Bayesian inference for the parameters of interest is based on the posterior distribution of w , θ , and σ^2 :

$$f(w, \theta, \sigma^2|y, \alpha, \beta, \gamma, K_{\min}, K_{\max}) \propto L(y|w, \theta, \sigma^2) f(w|\gamma) f(\theta|K_{\min}, K_{\max}) f(\sigma^2|\alpha, \beta) \quad (12)$$

Because of the complexity of (12), it is difficult to derive a closed-form formula for the posterior distribution, Teng and Liechty (2009). The Markov chain Monte Carlo (MCMC) simulation is therefore used to sample w , θ , and σ^2 . Because of the monotonicity of parameters w and θ in (4), an MCMC algorithm with Gibbs sampler can be used to avoid manual tuning procedures in the MCMC simulation.

2.3 Kernel smoothing density estimate of the quadrature method

The density $\hat{f}_N(x|w, \theta)$ from (3) is a weighted sum of δ functions and hence is not a continuous density. However, in many cases, it is interesting to visualize the SPD as a smoothed density. The kernel density for a set of M observed points $\vartheta = (\vartheta_1, \dots, \vartheta_M)^\top$ is:

$$\hat{f}(x|\vartheta) = \int \hat{g}(u)K_h(x-u)du = \frac{1}{M} \sum_{m=1}^M K_h(x - \vartheta_m) \quad (13)$$

where $K_h(\cdot) = h^{-1}K(\cdot/h)$ is a kernel function with a bandwidth h and $\hat{g}(u)$ the sum δ -functions

$$\hat{g}(u) = M^{-1} \sum_{m=1}^M \delta_{\vartheta_m}(u)$$

with locations ϑ . Obviously, different values of h will change the appearance of $\hat{f}(x|\vartheta)$. The Silverman's rule of thumb suggests a bandwidth

$$h_G = 1.06\hat{\sigma}M^{-1/5} \quad (14)$$

where $\hat{\sigma}$ is the sample standard deviation of ϑ and a normal kernel $K = \varphi$ the pdf of $N(0, 1)$, Silverman (1986).

Note that each ϑ_m for $m = 1, \dots, M$ appears with equal probability $1/M$. However, in the Bayesian quadrature method, θ_n appears with probability w_n , for $n = 1, \dots, N$. Therefore, we need to adjust the sample size and use $\hat{g}(u) = \sum_{n=1}^N w_n \delta_{\theta_n}(u)$ instead. The smoothed density version of (3) becomes

$$f_N^s(x|w, \theta) = \sum_{n=1}^N w_n K_h(x - \theta_n) \quad (15)$$

Ideally, if we can find an integer \hat{M} , so that there exist integers \hat{M}_n satisfying $\hat{M}_n/\hat{M} = w_n$ for $n = 1, \dots, N$. The adjusted sample size \hat{M} might be used when applying (14). This procedure however does not work in practice, because w_n would be an irrational number in most cases, and \hat{M} simply does not exist. An alternative is to round off each w_n to the q -th decimal, and set the adjusted sample size being 10^q . For example, we choose $q = 2$ in our study, and the smoothed SPD appears to be reasonable. In the i -th swipe of the MCMC algorithm, we obtain $w^{(i)}$ and $\theta^{(i)}$ and the smoothed SPD $f_N^s(x|w^{(i)}, \theta^{(i)})$. We then report the posterior mean and 95% credible regions of the smoothed SPD based on $\hat{f}(x|w^{(i)}, \theta^{(i)})$ point-wisely.

As a remark, the bandwidth can be adjusted to other kernels by a canonical bandwidth, Härdle et al. (2004). For example for the quartic kernel:

$$K(u) = \frac{15}{16}(1 - u^2)^2 \mathbf{I}\{|u| \leq 1\}, \quad (16)$$

the Silverman's rule of thumb h_G transforms into:

$$h_{QUA} = 2.62 \cdot h_G \quad (17)$$

The bandwidth may also be selected by cross validation or other methods. However, the bandwidth selection problem to obtain a smoothed SPD from a quadrature method for visualization is not the main contribution of this study and further discussion will be omitted.

3 Empirical analysis

We first introduce the weather derivatives (WD) market and present an overview on WD data. One major feature is the data sparsity, which makes most existing methods for estimating the SPD difficult. We then apply the described technique to estimate the SPD on WD data. Data sparsity really makes the estimation of a SPD a statistical challenge. Here, we will show the promising power of the Bayesian quadrature approach over other methods to estimate the SPD (lognormal, kernel density, mixture of log-normal).

3.1 Weather Derivatives

WDs are financial contracts designed to hedge weather risk. The most common contracts traded at CME are based on temperature indices $I(T_t)$, where T_t denotes the temperature at time t . These are the Heating Degree Days (HDD), the Cooling Degree Days (CDD), and the cumulative average temperature (CAT):

$$\begin{aligned} HDD(\tau_1, \tau_2) &= \sum_{t=\tau_1}^{\tau_2} \max(c - T_t, 0) \\ CDD(\tau_1, \tau_2) &= \sum_{t=\tau_1}^{\tau_2} \max(T_t - c, 0) \\ CAT(\tau_1, \tau_2) &= \sum_{t=\tau_1}^{\tau_2} T_t \end{aligned} \tag{18}$$

where c is a threshold (usually 65°F or 18°C) and $[\tau_1, \tau_2]$ with $\tau_1 < \tau_2$ is the temperature measurement period. The standard is that $[\tau_1, \tau_2]$ denotes a month of the year or as seasonal strips. The futures in question are delivering over a period $[\tau_1, \tau_2]$, and not at a fixed delivery time τ . The HDD index measures the demand for heating, while the CDD index measures the demand for cooling. Consequently, temperature indices are the underlying and not the temperature by itself.

Financial mathematical tools given in Benth et al. (2007), Benth et al. (2011) and Härde and López-Cabrera (2012) allow the pricing of the non-tradable underlying by risk adjusted conditional expectation. Hereby, the futures temperature price is given by:

$$F(t, \tau_1, \tau_2) = \mathbb{E}^Q [I(T_t) | \mathcal{F}_t] \tag{19}$$

where $\mathbb{E}^Q[\cdot]$ be any equivalent martingale measure and \mathcal{F}_t a filtration information set. Consequently, the European temperature call option price written on the futures price is

defined as:

$$C(K) = \exp\{-r\tau\} \int \max\{F(t, \tau_1, \tau_2) - K, 0\} f(x) dx \quad (20)$$

In order to compute (20), it is necessary to know the stochastic properties of the temperature process T_t under the “physical measure” measure and then adjust the risk measure, see Härdle and López-Cabrera (2012). Here we estimate the SPD, different to the afore mentioned approach, directly under the risk neutral measure. Note that (20) is exactly (5) for $f = f_N$.

The options at CME are cash settled i.e. the owner of a future receives 20 times the Degree Day Index at the end of the measurement period, in return for a fixed price. At time t , CME trades contracts with different measurement periods $t \leq \tau_1 < \tau_2$ or different maturities $\tau = \tau_2 - t$. The measurement period for CAT/HDD futures is typically during April-November, while CDD futures are measured during November-April.

3.2 Overview on the WD Data

The WD data was purchased from CME for the study period from 2002/01/02 to 2012/05/11. The reported price is the settlement price for the future or option contract, and the volume is the number of contracts traded.

Depending on the measurement period, temperature products in the US market are further categorized into monthly, seasonal strips, and average products. HDD monthly products have seven contract months: October, November, December, January, February, March, April, and CDD monthly products have seven contract months: April, May, June, July, August, September, and October. For HDD seasonal strips, the contract period covers from October to April, and for CDD seasonal strips, the contract period covers from April to October. Contract for weekly average products covers all five weeks. Table 1 gives an overview of the volume of the temperature market.

Figure 2 illustrates the volume for US temperature futures and options in the study period. The trading activity increased dramatically since 2002 but declined after the 2008 financial crisis. This is surprising since one could expect that these markets are uncorrelated with financial markets. However we believe that the decline is due to ignorance of the temperature market as a intermediary for diversification purposes. Star plots in Figure 3 divide the volume into HDD-CDD monthly, HDD-CDD seasonal strips, and weekly average for futures and options for each US city. A star plot represents each city as a star whose i -th spoke is proportional in length to the volume size of i -th product (HDD Monthly, CDD Monthly, HDD Strips, CDD strips, Average) of the observed city. Clearly, monthly products are the most popular traded products, followed by seasonal strips. Nevertheless, no weekly average products are really traded in the US temperature market.

The volume of HDD, CDD, Average monthly and seasonal strips futures and options for all US cities is reported in Table (1. New York is to be the biggest temperature market and takes about 20% of the market volume, followed by Chicago (10%), Atlanta (9%), Cincinnati (8%), and Dallas (7%). The market share of these five cities exceeds 50% of the

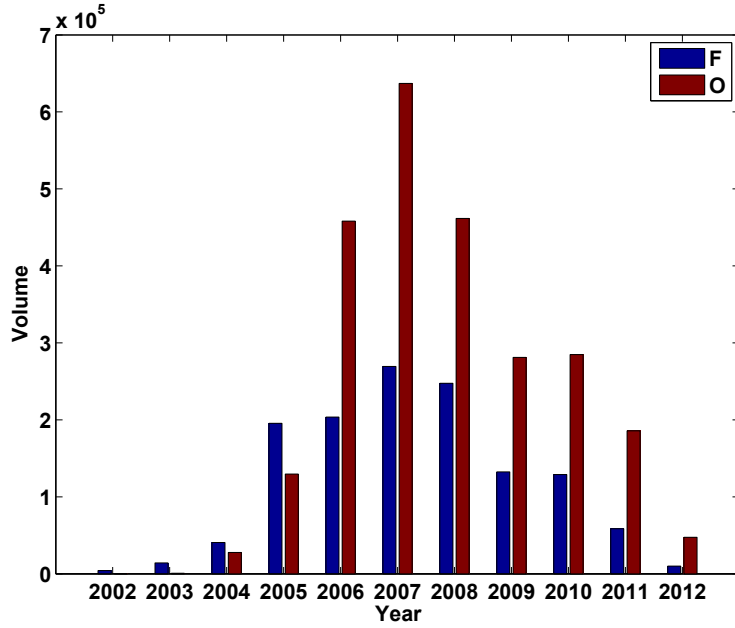


Figure 2: The volume for US temperature futures (F) and options (O).

US temperature market. Following this, we took these cities as the most representative cities. Figure 4 gives time series plots for New York, Atlanta, Chicago and Dallas monthly HDD, CDD monthly and seasonal strip futures prices. The futures market is more liquid but also more volatile than option prices. In addition, most HDD and CDD futures are traded only with time to maturities less than a year. These features of future prices make the pricing mechanism for weather derivatives unique and challenging.

We further divide the volume of HDD and CDD monthly and HDD seasonal strip options with respect to strike prices and time to maturities, as summarized in Table 2. It is shown that most options are traded with only a few number of strike prices and of a short time to maturity (within one month and less than a year). Because of the fact that options are only traded with a few number of strike prices, this data sparsity problem makes most existing nonparametric methods (such as mixture of lognormal models or kernel methods) very difficult.

3.3 Illustration and fit of the technique

As depicted in Figure 5, we calibrate the SPD for HDD-CDD monthly and Seasonal strip options. These four plots present a typical pattern of option prices of weather options: options were traded only with a very few number of strike prices, sometimes only call options or put options were traded or the both of them. There is a trade-off in the selection of N . When N is larger, one produces better fit because there exist more free parameters in the model, but drawbacks of model complexity and computational demanding come along with. Because options are traded with only a few number of strike prices, we select $N = 5$ based on the overall availability of data.

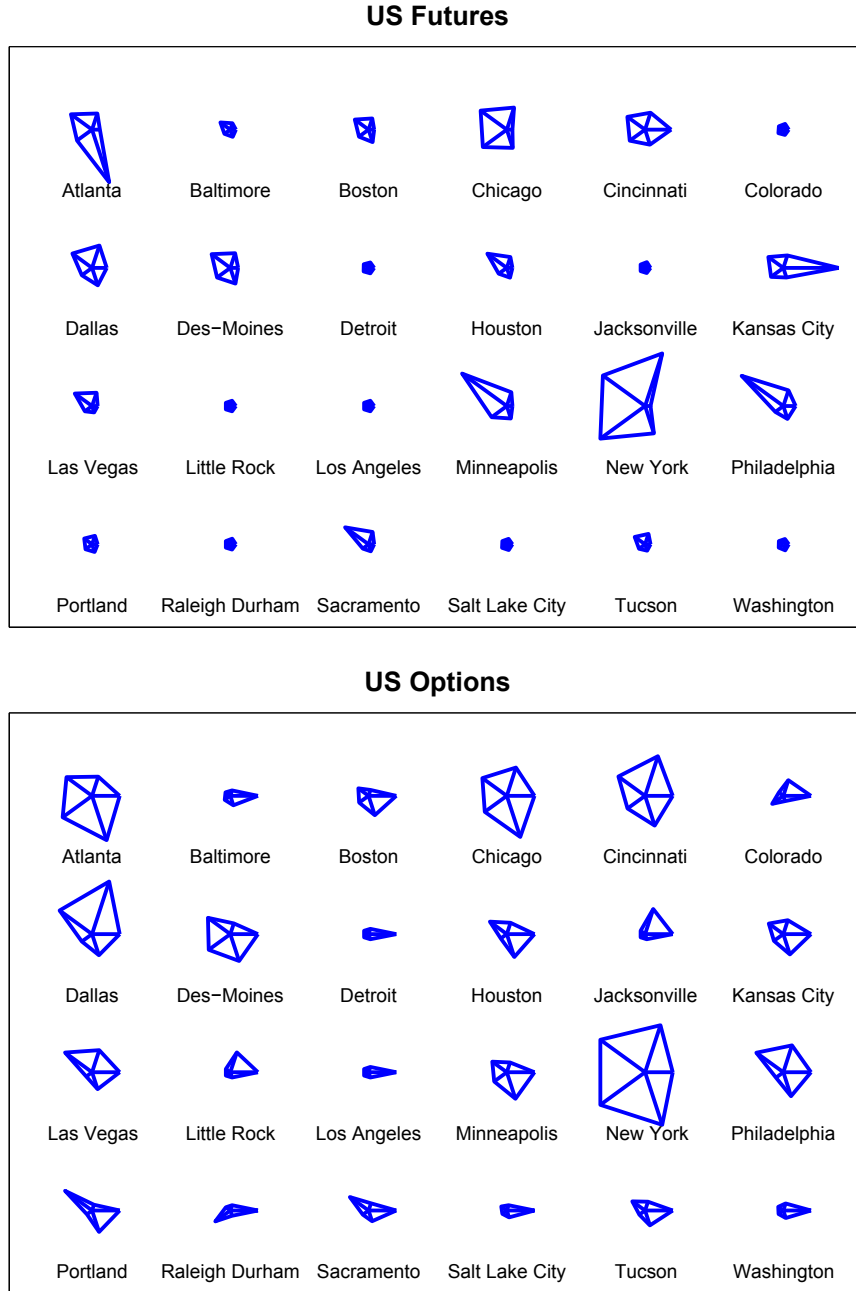


Figure 3: Star plots representing the volume for US temperature contracts (HDD-CDD monthly, HDD-CDD seasonal strips, and weekly average) futures (upper panel) and options (lower panel) for each city. Each city is represented as a star whose i -th spoke is proportional in length to the volume size of i -th product (HDD Monthly, CDD Monthly, HDD Strips, CDD strips, Average) of the observed city.

Table 1: The volume for HDD-CDD monthly, seasonal strips and average temperature products in each US city.

Index	City	Future				Option				Strips	CDD	Avg	MS (%)	Rank
		HDD		CDD		HDD		CDD						
		Monthly	Strips	Monthly	Strips	Monthly	Strips	Monthly	Strips					
1	Atlanta	49621	35567	14400	2150	50	56431	11647	117165	71950	0	9.44	3	
2	Baltimore	6633	3545	600	700	0	2600	100	12500	1100	0	0.73	16	
3	Boston	24178	19066	2200	1150	0	11029	550	42174	19450	0	3.15	13	
4	Chicago	90585	54950	3975	2800	0	39676	19300	107616	67725	0	10.17	2	
5	Cincinnati	50155	38035	2967	1700	455	29280	28910	73255	74975	0	7.89	4	
6	Colorado Springs	1936	1450	0	0	0	15025	8750	0	0	0	0.71	17	
7	Dallas	27206	55540	3700	1961	200	13085	39775	47450	94850	0	7.47	5	
8	Des Moines	40929	30510	3190	1450	50	38631	4460	64790	60900	0	6.44	6	
9	Detroit	2185	351	50	50	0	0	0	0	0	0	0.07	23	
10	Houston	16901	18229	1400	1700	0	3700	5000	52950	33950	0	3.52	12	
11	Jacksonville	100	1600	0	0	0	0	16575	0	0	0	0.48	19	
12	Kansas City	36513	23145	1325	1350	1100	11025	7200	45050	33750	0	4.22	10	
13	Las Vegas	12680	26635	325	1650	0	3100	14200	34600	76650	0	4.47	9	
14	Little Rock	120	105	0	0	0	0	12250	0	0	0	0.33	20	
15	Los Angeles	100	400	0	0	0	0	50	0	0	0	0.01	24	
16	Minneapolis	50085	27955	2150	1500	0	18206	3850	63350	34000	0	5.29	8	
17	New York	187264	154605	6700	4860	0	90620	35175	141850	136350	0	19.93	1	
18	Philadelphia	16441	34449	2300	2250	150	6408	18210	56000	76150	0	5.59	7	
19	Portland	10329	10855	725	450	0	1720	450	48200	76450	0	3.92	11	
20	Raleigh Durham	550	1500	0	0	0	23700	0	0	0	0	0.68	18	
21	Sacramento	6383	23401	550	750	0	2850	1675	16200	48000	0	2.63	14	
22	Salt Lake City	739	504	150	0	0	0	0	4500	3500	0	0.25	22	
23	Tucson	7283	16965	350	750	0	2700	3010	28750	27800	0	2.30	15	
24	Washington	550	25	250	0	0	650	1500	6650	2000	0	0.31	21	

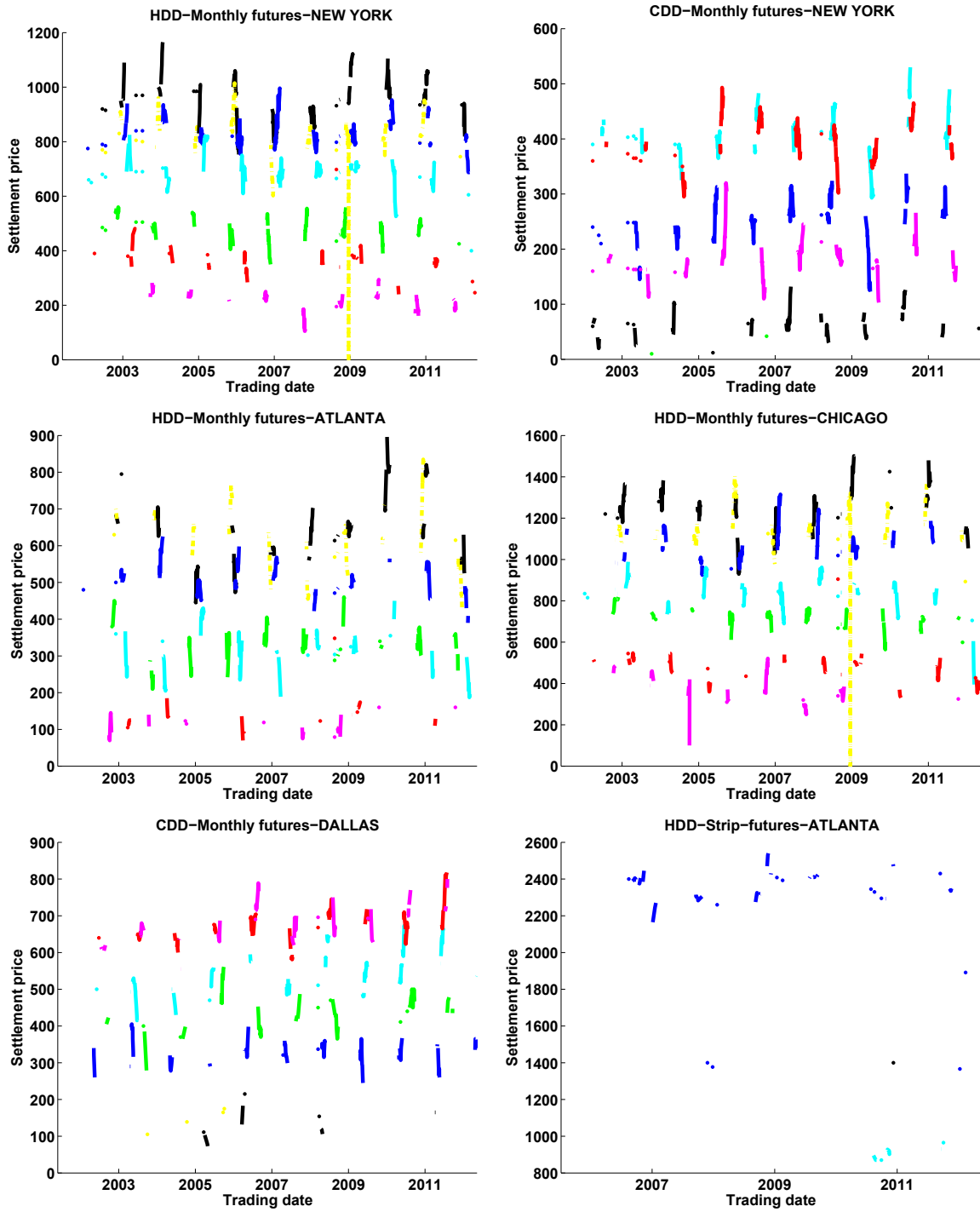


Figure 4: Time series plots of New York, Atlanta, Chicago, Dallas HDD/CDD monthly and seasonal strips futures prices. HDD monthly futures with the measurement period of January (Black), February (Blue), March (Cyan), September (Red), October (Magenta), November (Yellow) and December (Green). CDD monthly futures with the measurement period of May (Black), June (Blue), July (Cyan), August (Red), and September (Magenta). HDD seasonal strip with the measurement period in January (Blue), March (Cyan) and December (Cyan)

Table 2: The number of transactions - trading days (TD) and volume (vol) of New York/ Atlanta/Chicago/Dallas HDD and CDD monthly and HDD seasonal strip options with respect to time to maturity (τ) in month and the number of strike prices.

HDD - New York										HDD - Atlanta					HDD - Chicago				
τ		Number of strike prices																	
		1	2	3	4	Total	1	2	3	4	5	Total	1	2	3	4	Total		
≤ 1	TD	71	23	7	1	102	56	12	4	1	1	74	50	10	-	-	60		
	vol	17495	12650	9900	1400	41445	12861	4700	2950	700	1250	22461	10961	4975	-	-	15936		
(1,2]	TD	54	26	3	4	87	39	26	2	1	-	68	32	13	2	2	49		
	vol	12450	21700	1075	5400	40625	10245	19825	2800	1000	-	33870	50	2000	-	-	2050		
(2,3]	TD	3	1	-	-	4	1	-	-	-	-	1	2	-	-	-	2		
	vol	1000	1000	-	-	2000	100	-	-	-	-	100	2000	-	-	-	2000		
(3,4]	TD	2	1	-	-	3	-	-	-	-	-	-	1	-	-	-	1		
	vol	300	2000	-	-	2300	-	-	-	-	-	-	2000	-	-	-	2000		
(4,5]	TD	1	1	-	-	2	-	-	-	-	-	-	1	-	-	-	1		
	vol	250	2000	-	-	2250	-	-	-	-	-	-	2000	-	-	-	2000		
(5,6]	TD	-	1	-	-	1	-	-	-	-	-	-	1	-	-	-	1		
	vol	-	2000	-	-	2000	-	-	-	-	-	-	2000	-	-	-	2000		

CDD - New York										HDDstrips - Atlanta										CDD- Dallas			
τ		Number of strike prices																					
		1	2	3	4	Total	1	2	3	4	5	6	7	8	Total	1	2	3	4	Total			
≤ 1	TD	43	3	1	-	47	1	6	-	6	-	-	-	-	13	40	-	90	1	131			
	vol	17425	2000	600	-	20025	200	3100	-	2700	-	-	-	-	6000	12400	9150	1250	-	-			
(1,2]	TD	34	13	2	-	49	6	6	1	-	-	-	-	-	13	30	-	10	1	41			
	vol	8200	5750	1200	-	15150	4700	6250	1875	-	-	-	-	-	-	9125	5300	750	-	-			
(2,3]	TD	-	-	-	-	-	3	9	-	3	-	-	-	-	15	-	1	-	-	1			
	vol	-	-	-	-	-	2240	7500	-	5500	-	-	-	-	-	-	450	-	-	450			
(3,4]	TD	-	-	-	-	-	-	9	-	-	-	-	-	-	9	-	1	-	-	1			
	vol	-	-	-	-	-	-	10400	-	-	-	-	-	-	10400	-	450	-	-	450			
(4,5]	TD	-	-	-	-	-	1	11	1	4	1	1	1	-	19	-	1	-	-	1			
	vol	-	-	-	-	-	1750	10500	1000	9500	-	6500	6500	-	-	-	450	-	-	450			
(5,6]	TD	-	-	-	-	-	1	11	-	3	-	-	-	-	16	-	-	-	-	-			
	vol	-	-	-	-	-	250	11700	-	6750	-	-	-	1	-	-	-	-	-	-			
(6,7]	TD	-	-	-	-	-	10	-	-	-	-	-	-	-	1	-	-	-	-	-			
	vol	-	-	-	-	-	9000	-	-	-	-	-	-	-	9000	-	-	-	-	-			
(7,8]	TD	-	-	-	-	-	3	-	-	-	-	-	-	-	3	-	-	-	-	-			
	vol	-	-	-	-	-	6000	-	-	-	-	-	-	-	6000	-	-	-	-	-			

To calibrate (5) and (6), we implement an MCMC algorithm to explore the posterior distribution in (12). Trace plots of the log-likelihood in Figure 6 show that the MCMC algorithm is converging. The traces indicate that the MCMC approaches a stationary distribution. Figure 5 imposes model prices of the last swipe of the MCMC algorithm and demonstrates the fit, because model prices are close to market prices. For residual analysis, we calculate the residual at each swipe of the MCMC algorithm by

$$r_{ijk} = \log y_{ijk} - \log C_{ij}^N(w, \theta) \quad (21)$$

and provide kernel smooth density plots of the posterior distribution of these residuals in Figure 7. All these four panel plots demonstrate that the residuals have mean zero, and are symmetric about zero when comparing with the normal KDE. This visual presentation supports our error assumption as a normal distribution in (8).

Figure 8 presents trace plots, autocorrelation plots, and KDE plots for σ^2 in the MCMC algorithm. These trace plots show that the MCMC simulation appears to be convergent. Autocorrelation plots show that the MCMC samples are efficient. The KDE plots give a visualization of the posterior distribution of σ^2 and suggest that the inverse gamma is a good candidate for a posterior.

The density (3) approximates the SPD by weighted sum of δ functions and is discontinuous by its nature. As described earlier, to produce a smoothed SPD for visualization, we round off each w_n to the second decimal, and set the adjusted sample size as 100. Then we employ the kernel density estimation with a Gaussian kernel $K(\cdot) = \varphi(\cdot)$ and a bandwidth selected using the rule of thumb in (14) to calculate a smoothed SPD at each swipe in the MCMC algorithm.

Thus, it is clear that the smoothed density version (15) becomes:

$$\begin{aligned} f_N^s(x|w, \theta) &= \sum_{n=1}^N w_n K_h(x - \theta_n) \\ &= \sum_{n=1}^N w_n \frac{1}{h} \varphi\left(\frac{x - \theta_n}{h}\right) \\ &= \sum_{n=1}^N w_n \varphi(x; \theta_n, h) \end{aligned}$$

where $\varphi(x; \theta_n, h)$ is the pdf of $N(\theta_n, h^2)$ distribution.

Collecting these smoothed SPD, Figure 9 gives the posterior mean (red line) and 95% credible regions (blue dotted lines) of the implied SPD. The right-upper and left-lower pictures show that the 95% credible regions are tight to the posterior mean of the smoothed SPD, whereas the other pictures depict that the 95% credible regions are wide. The cluster of star points in the horizontal axis denote the future prices.

In Bayesian analysis, the 95% credible region for the smoothed SPD provides a region where 95% of the posterior distribution of the smoothed SPD will fall into. In the case of HDD New York options with maturity in 2 months traded at 20050121 and the case of CDD Dallas Option one month to maturity traded at 20100604, the left tail of 95%

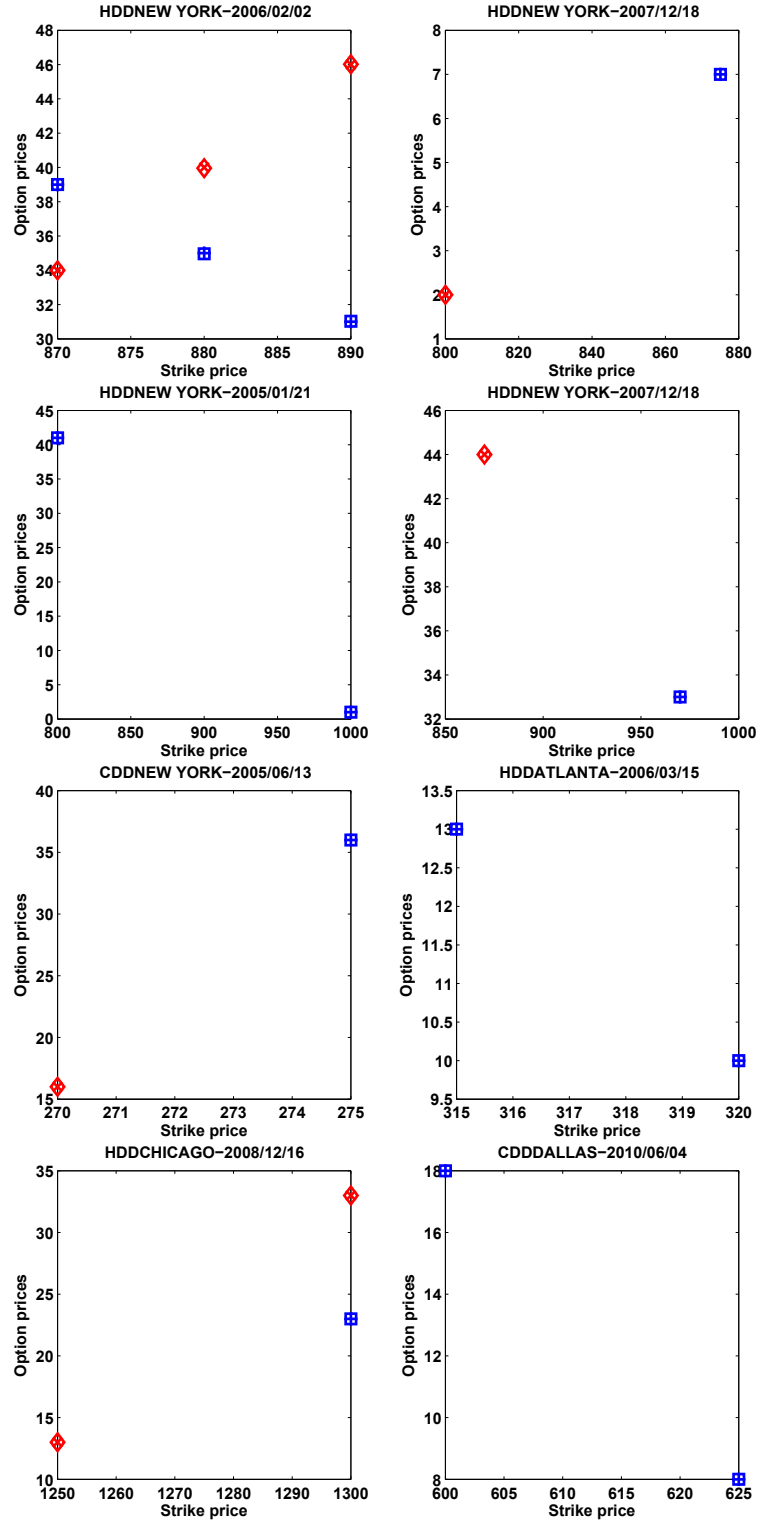


Figure 5: Plots of market prices of New York/Atlanta/Chicago/Dallas HDD-CDD monthly options and model prices of the last swipe in the MCMC algorithm using a quadrature method with $N = 5$. For market prices, a call option is indicated with a blue plus and a put is indicated with a red cross. For model prices, a call option is indicated with a blue diamond, and a put option is indicated with a red square.

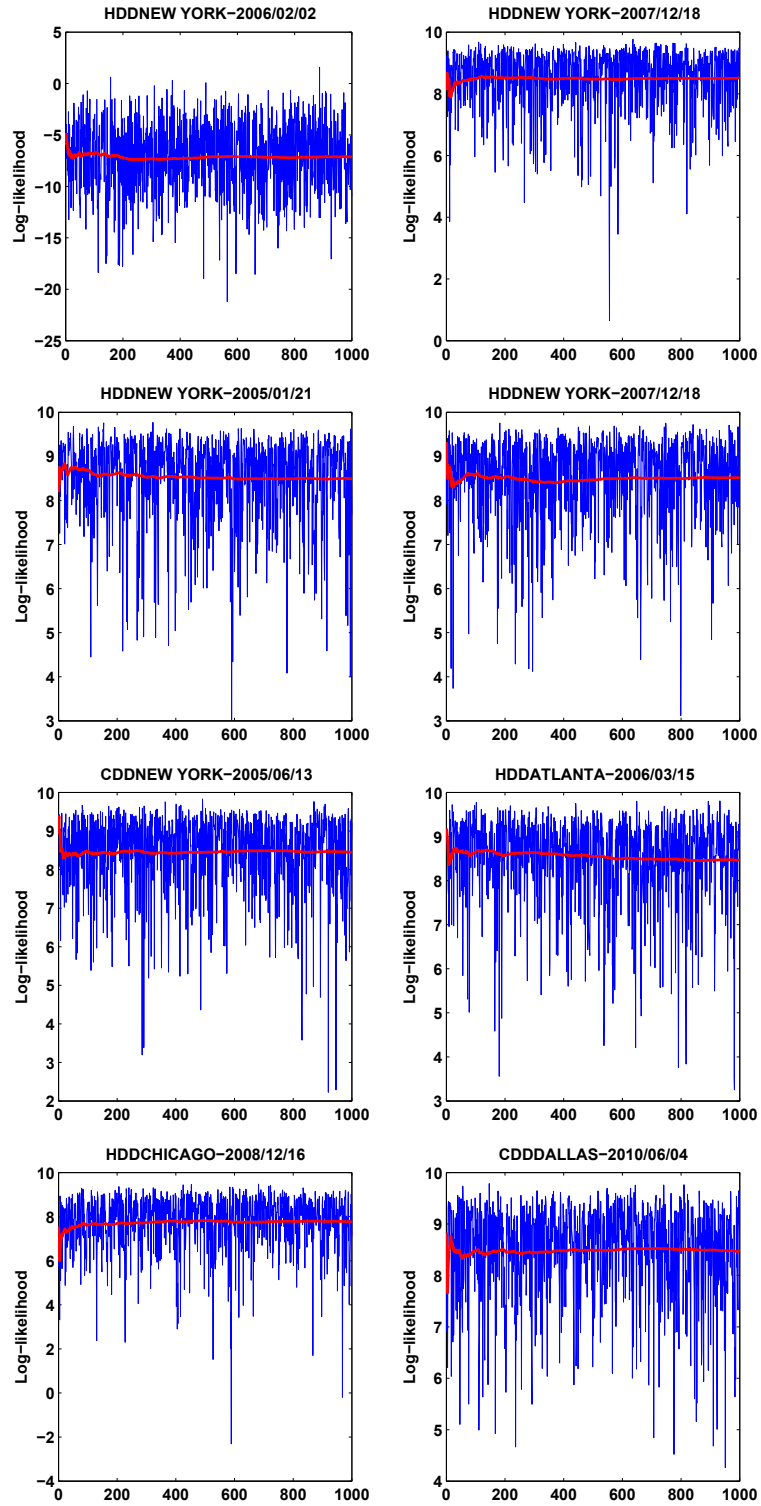


Figure 6: Trace plots for the log-likelihood (blue line) and mean (red line) of the MCMC algorithm for different trading dates.

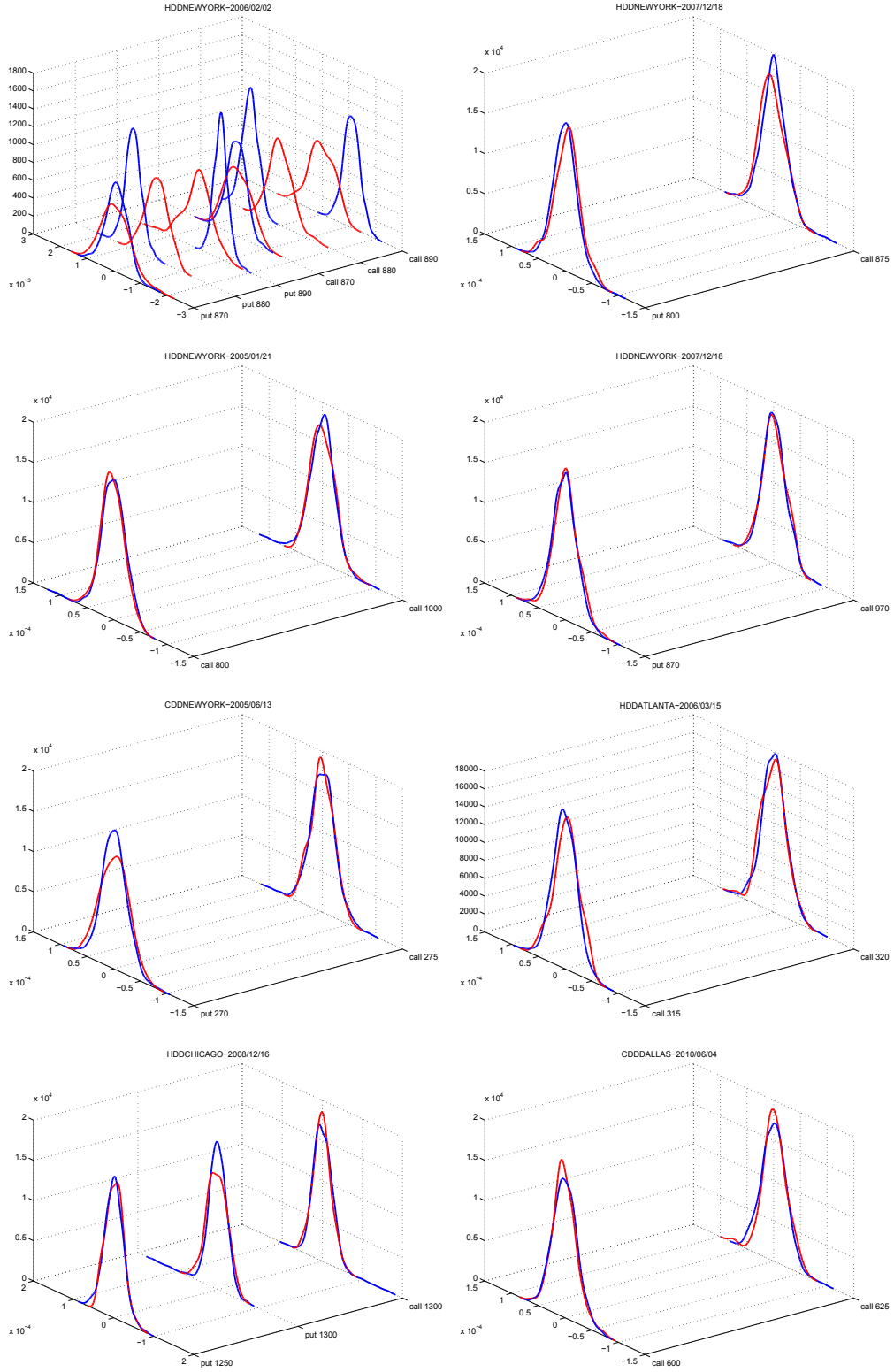


Figure 7: Kernel density estimate (KDE) plots for the posterior distribution of residuals for HDDs and CDDs products (in blue) vs the normal KDE (red).

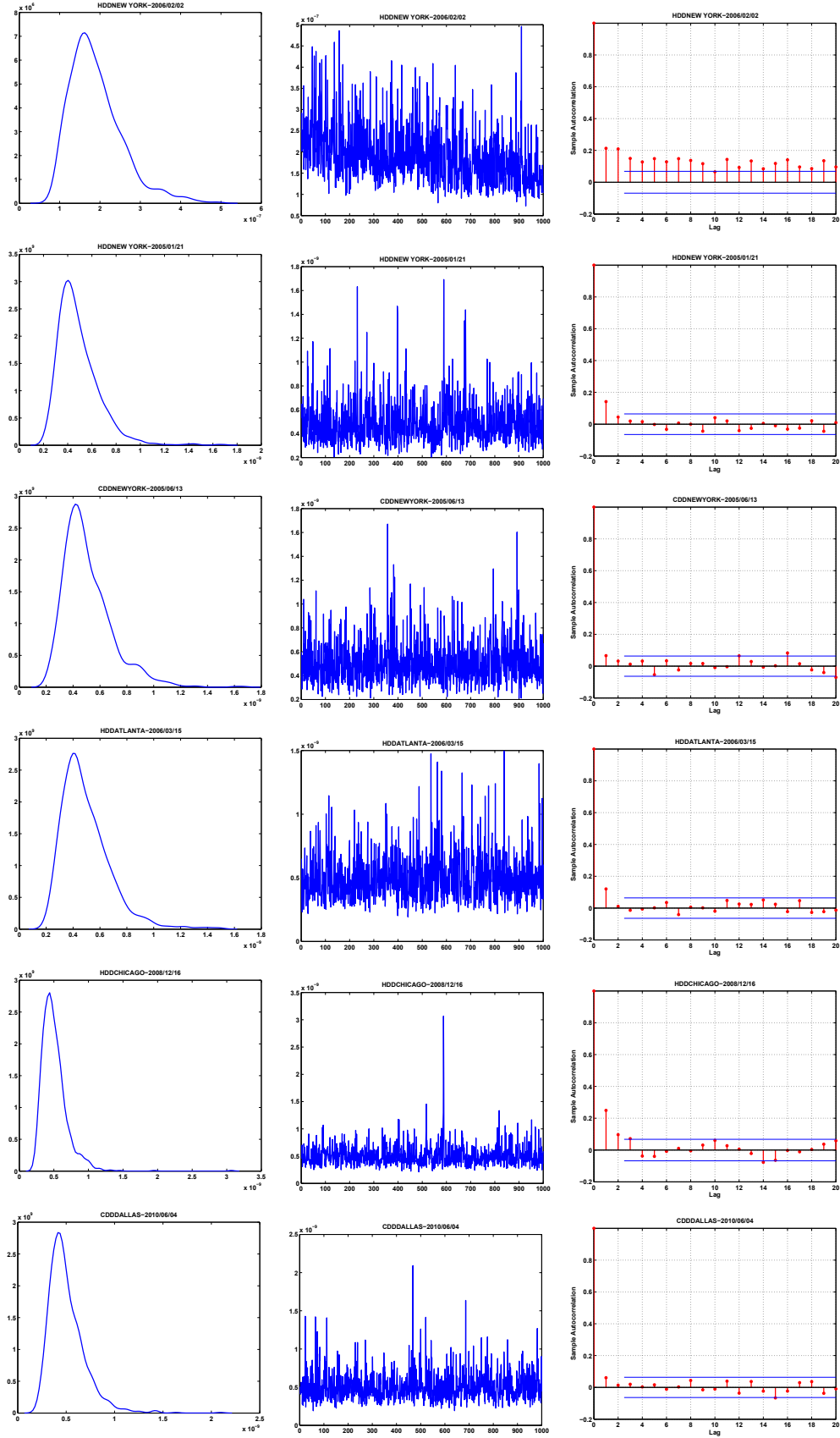


Figure 8: Kernel density estimate, trace plots and autocorrelation plots of σ^2 of the MCMC simulation.

credible region appears to be extremely wide. This feature is not surprising though, because the data set for calibration consists of call options with only two strike prices, namely, 800 and 1000 and 600 and 625 respectively. Indeed, a call option price is simply the expected future price distracting the strike price under the SPD for the future price larger than the strike price, and multiplied by a discounted factor. Therefore, an option price only provides information for the right tail of the SPD. Once a few quadrature points in the right tail have achieved a high likelihood, points of the quadrature in the left tail (in this case, smaller than 800) do not affect the likelihood. As a result, these points are influenced only by its prior distributions. The prior assumptions in (9) and (10) put simply vague information for the weights and locations in the quadrature method. Such an assumption allows points in the left tail of the quadrature method moving freely, and causes a wider credible region in the left tail, as demonstrated in the left-upper panel in Figure 9.

Similarly, for the right-lower panel, the 95% credible region is wider around strike 1000 but is tight in two side tails. This is because the data set consists of one call with strike 970 and one put with strike 870. As a result, the call option price gives information of the right tail of the SPD, whereas the put option price gives information of the left tail of the SPD. When some points of both right and left tails in the quadrature method have achieved a high likelihood, points of the quadrature around 920 would not affect the likelihood. These points are determined by their prior assumptions again, and provide a wider credible region around 920.

Selecting prior distributions for the quadrature method is critical. In this research, we choose vague prior assumptions for the parameters and the analysis successfully reveals the fact that the width of the 95% credible region depends highly on the information provided by option prices and the prior assumptions on the parameters in the quadrature method. One may adopt more sophisticated prior distributions based on experience and knowledge. This flexibility may be considered as a technical advantage of the Bayesian quadrature method.

3.4 The equivalence/infeasibility to other nonparametric methods

To illustrate the feasibility of the proposed methodology, we compare with other very known approaches. Recall that nonparametric methods are popular tools avoiding risk of misspecification. Some of these methods estimate the SPD by differentiating an interpolation of smoothing of option prices. In this context, data sparsity makes the estimation of the SPD a statistical challenge. Let us now explain why the kernel regression method and also mixtures of lognormals do not work well in the context of WD implied SPDs.

The cross-section of the call and put prices are given in Figure 11. Different grids correspond to different contracts with different times to measurement periods and consequently one can argue that option prices can be extrapolated as a smoothed function of the strike.

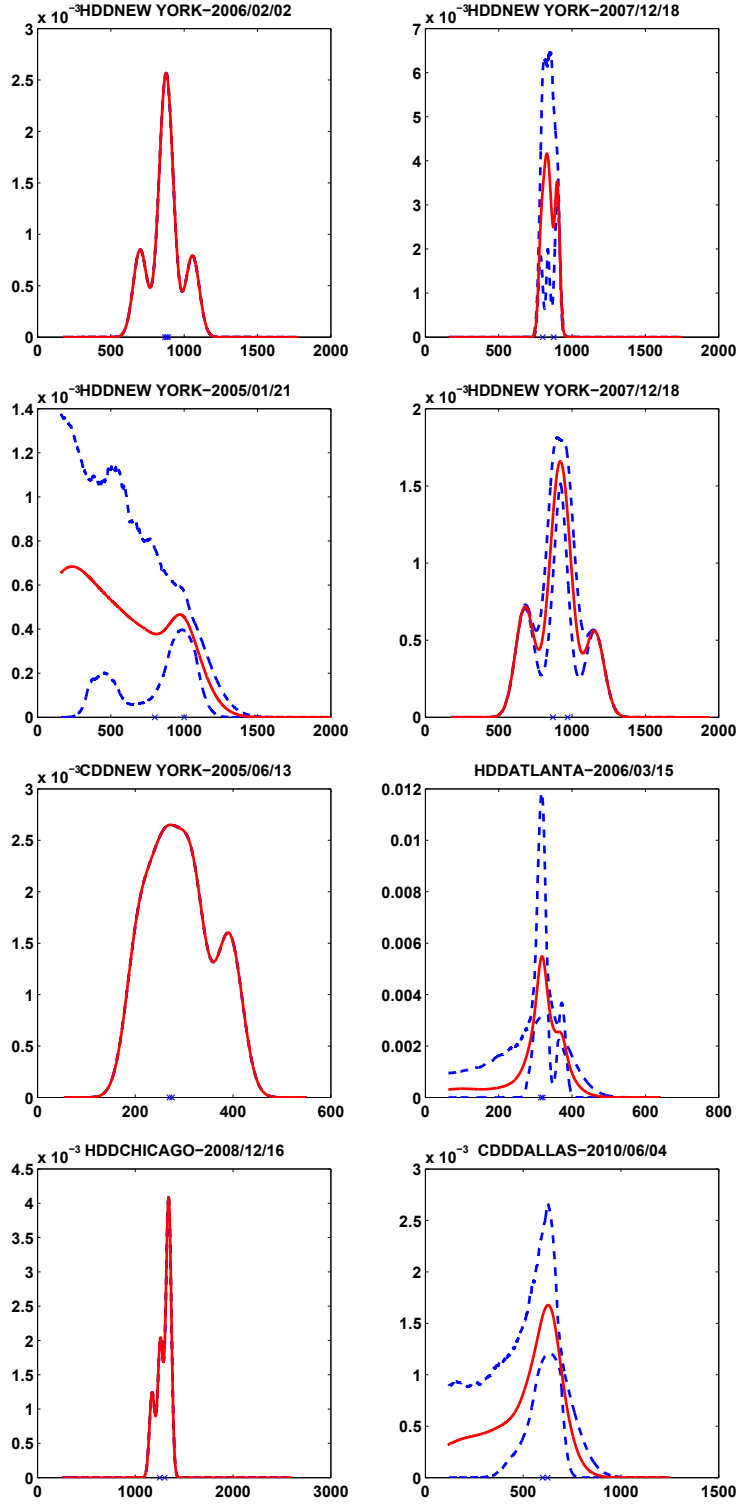


Figure 9: The posterior mean (in a red solid line) and the 95% credible region (in blue dashed lines) of the smoothed SPD implied from New York HDD monthly options with respect to trading dates and time to maturity. The cluster of star points in the horizontal axis denote the future prices.

3.4.1 Kernel Regression

The kernel regression method (KRM) takes advantage of differentiating twice (1) :

$$f(K) = e^{r\tau} \frac{\partial^2}{\partial K^2} C(K) \quad (22)$$

In order to employ (22) one needs more observations and therefore relies on the put-call parity to transfer put option prices to call option prices. When the market is illiquid, it is not promising to employ the put-call parity though. In our empirical data analysis most options are traded with only a few strike prices. Very often, an option was traded only with one or two strike prices. When the kernel method is applied to a dataset of such a case, it is even difficult to find an option function $C(K)$, not to mention to find its second derivatives.

For illustration, we consider SPD estimation using the kernel regression method for New York HDD monthly options. As we observe, the traditional idea of estimating the SPD by differentiating an interpolation of option prices does not hold in practice. Furthermore, there is no guarantee that the estimated SPD is positive as required by any density. Consequently, KRM is sensitive to data sparsity.

3.4.2 Mixture of lognormals

When applying the mixture of lognormal methods, it is necessary to specify the range of the variances of the log-normal density. The selection is objective and influences the estimated SPD dramatically. When the dataset consists of a few data point, it is possible to produce two totally different densities (particularly in terms of variances) which produce the same quality of model fit.

We estimate the SPD using a mixture of lognormal for New York HDD monthly options, given in Figure 5. For mixture of lognormals, we will show that two different SPD using mixture of lognormal produce the same model fit, but they have quite different higher moments.

Yuan (2009) proposed a function class:

$$\mathcal{F} = \left\{ f(\cdot) : f(x) = \int f(x|\mu, \sigma^2) dG(\mu, \sigma), \text{supp}(g) \subset [-M, M] \times [\underline{\sigma}, \bar{\sigma}] \right\}$$

where $M < \infty$ and $0 < \underline{\sigma} \leq \bar{\sigma} < \infty$, $f(x|\mu, \sigma^2)$ is the pdf of the lognormal distribution with location μ and scale σ and G determines the mixing distribution. The corresponding pricing function in this case is similar to (2):

$$\begin{aligned} C(X; \mu, \sigma^2) &= \exp(-r\tau) \int_0^\infty \wp_{ij}(x) f(x) dx \\ C(X; G) &= \int C(X; \mu, \sigma^2) dG(\mu, \sigma) \end{aligned} \quad (23)$$

as the SPD $f(x)$ is defined in the previous family \mathcal{F} . The least squares estimate of the pricing function can be written as:

$$\hat{G}(\cdot) = \arg \min_{G \in \mathcal{G}} n^{-1} \sum_{i=1}^n \{y_{ijk} - C(X; G)\}^2 \quad (24)$$

where \mathcal{G} is the collection of all probabilities measures on μ and σ^2 . Note that the minimization is taken over a function space of infinite dimensions, however the solution can be represented in a finite dimensional space. In particular, all solutions can be expressed as a convex combination of at most $n + 1$ Black Scholes type of pricing functions.

This model has several nice theoretical properties. For example, as the sample size n increases, the pricing functions can be recovered with squared error converging to zero at the rate of $\log^2 n/n$, which is close to the parametric rate of convergence $1/n$. However, practical difficulties arise when fitting mixtures of lognormal distributions (or other mixtures models) to real data. The feature that weather options are traded with a few number of strike prices make mixture models inapplicable, because mixture models need to select corresponding scale parameters and the number of components. For example, when options are traded with n different strike prices, maximum likelihood suggests to use $n/2$ support points. When n is large, this leads a very complicated model and possible over fitting problem. When n is small, the resulting model may be inappropriate. In addition, numerical procedure for searching the maximum likelihood estimate is particularly difficult for large n .

We apply mixture of lognormals by Yuan (2009) to the New York monthly HDD call options traded on 2006/02/02, with two different manually selected variances. Figure 10 shows that these two estimated SPD are quite different in shapes, although they produce similar quality of model fit. Therefore, this illustration shows that the estimated SPD is very sensitive to the selection of $\underline{\sigma}$. In practical implementation, Yuan (2009) suggests to determine $\underline{\sigma}$ by cross-validation. This however is very computationally demanding.

3.5 Dynamics of SPD

Table 3 records the number of trading days with respect to the number of strike prices and trading months for the New York HDD/CDD monthly options with time to maturity τ less than one month and Atlanta HDD Seasonal Strips with $\tau = 6$. Again, the data sparsity remains an issue in the biggest weather market, the New York market. For illustration, we implement (4) for every trading day in March 2006. For other months, most of the number of strike in each trading day is simply one, and such case makes the SPD estimation very difficult in the sense that option price only provides information for one side of the SPD.

Figure 11 plots give the evolution of New York-HDD, Atlanta HDD-Seasonal and option prices with time to maturities in one and six months respectively, against strike prices and trading days in March 2006 and October 2007. It is clear that options were traded with very few strike prices (from one to four strike prices) during this month. We used five support points ($N=5$) in the quadrature method, and calculate R^2 in a logarithmic

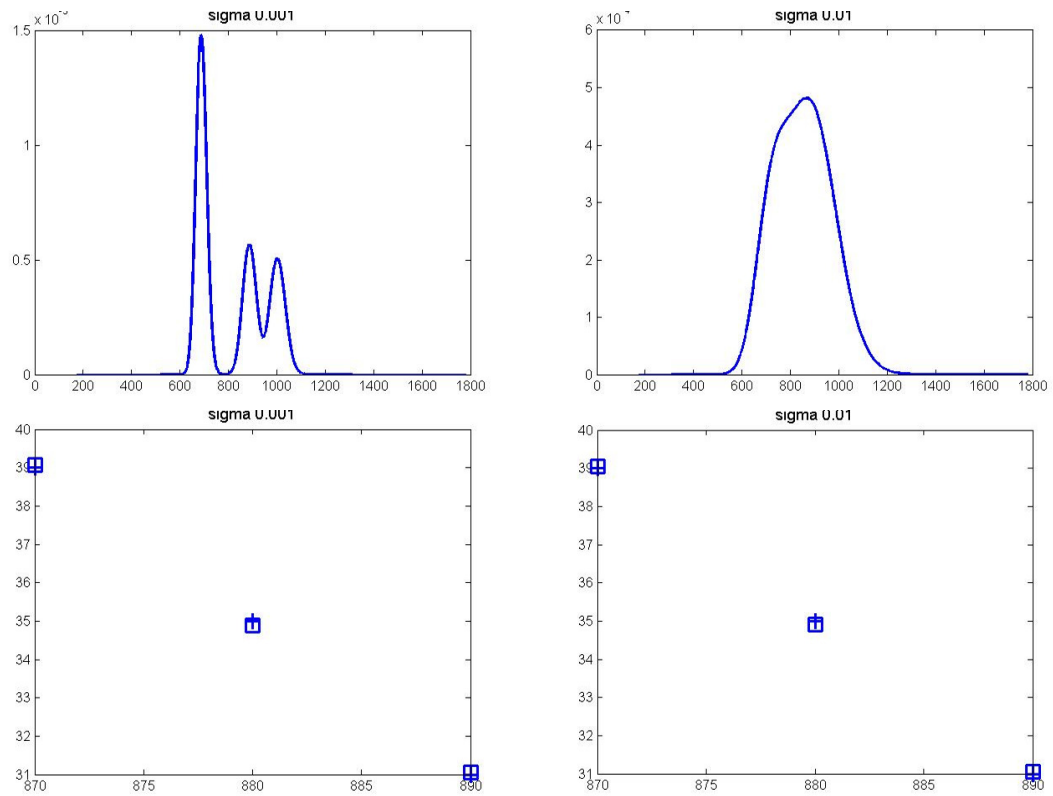
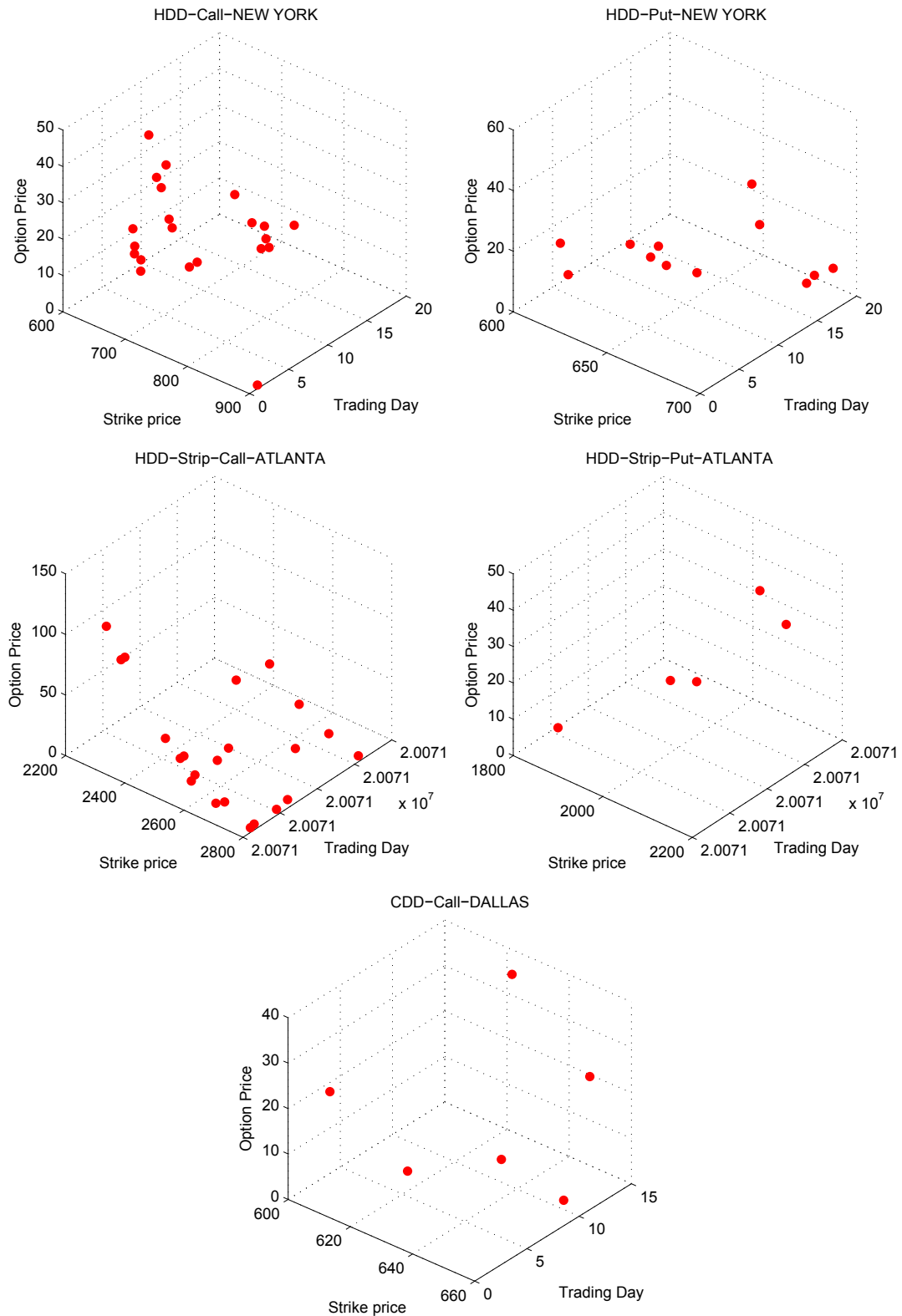


Figure 10: Estimated SPD using mixtures of lognormal. The left-upper and right-upper panels give estimated SPD using σ equal to 0.01 and 0.001, respectively. The left-lower and right-lower panels give market prices indicated by a blue cross and model prices indicated by a blue square.

Table 3: Number of trading days with respect to the trading month and the number of strike prices for HDD/CDD monthly options with time to maturity less τ than one month and Atlanta HDD Seasonal Strips with $\tau = 6$ (time of measurement period of 5 months).

Nr. of strike prices											Nr. of strike prices												
Year	Type	City	Month	τ	1	2	3	4	5	Total	Year	Type	City	Month	τ	1	2	3	4	5	6	Total	
2002	HDD	NY	11	1	1	-	-	-	-	1	2006	HDD-Strip	Atlanta	10	6	-	-	1					1
2004	HDD	NY	1	1	2	-	-	-	-	2	2007	HDD-Strip	Atlanta	10	6	-	-	3	1	2	1	1	8
2004	HDD	NY	2	1	1	-	-	-	-	1	2008	HDD-Strip	Atlanta	10	6	-	-	4		1	1	1	5
2004	HDD	NY	3	1	1	-	-	-	-	1	2009	HDD-Strip	Atlanta	10	6	-	-	1	1		1		3
2005	HDD	NY	1	1	1	-	-	-	-	1	2010	HDD-Strip	Atlanta	10	6	-	-	2					2
2005	HDD	NY	2	1	2	-	-	-	-	2	2002	HDD	Chicago	11	1	1	-						1
2005	HDD	NY	3	1	1	-	-	-	-	1	2002	HDD	Chicago	12	1	2	-						2
2005	HDD	NY	12	1	1	1	-	-	-	2	2004	HDD	Chicago	2	1	1	-						1
2006	HDD	NY	1	1	2	-	-	-	-	2	2005	HDD	Chicago	2	1	2	-						2
2006	HDD	NY	2	1	5	3	1	-	-	9	2005	HDD	Chicago	12	1	4	-	-					4
2006	HDD	NY	3	1	3	5	4	1	-	13	2006	HDD	Chicago	3	1	5	2						7
2006	HDD	NY	10	1	1	1	-	-	-	2	2006	HDD	Chicago	10	1	2	1						3
2006	HDD	NY	11	1	1	2	-	-	-	3	2006	HDD	Chicago	11	1	2							2
2006	HDD	NY	12	1		1	-	-	-	1	2007	HDD	Chicago	1	1		3						3
2007	HDD	NY	1	1	2	-	-	-	-	2	2007	HDD	Chicago	2	1		1						1
2007	HDD	NY	2	1	4	-	-	-	-	4	2007	HDD	Chicago	3	1	1							1
2007	HDD	NY	3	1	1	-	-	-	-	1	2007	HDD	Chicago	12	1		1						1
2007	HDD	NY	11	1	1	-	-	-	-	1	2008	HDD	Chicago	1	1		1						1
2007	HDD	NY	12	1	1	2	-	-	-	3	2008	HDD	Chicago	2	1	3							3
2008	HDD	NY	1	1	6	1	2	-	-	9	2008	HDD	Chicago	3	1	2	1						3
2008	HDD	NY	2	1	3	-	-	-	-	3	2008	HDD	Chicago	12	1	2	2						4
2008	HDD	NY	12	1	2	1	-	-	-	3	2009	HDD	Chicago	1	1	2							2
2009	HDD	NY	1	1	4	-	-	-	-	4	2009	HDD	Chicago	12	1	1							1
2009	HDD	NY	2	1		2	-	-	-	2	2010	HDD	Chicago	3	1	1							1
2009	HDD	NY	3	1		1	-	-	-	1	2010	HDD	Chicago	11	1	1							1
2009	HDD	NY	11	1	3	-	-	-	-	3	2010	HDD	Chicago	12	1	2							2
2009	HDD	NY	12	1		1	-	-	-	1	2011	HDD	Chicago	1	1	3							3
2010	HDD	NY	3	1	1	-	-	-	-	1	2011	HDD	Chicago	2	1	2							2
2010	HDD	NY	11	1	5	-	-	-	-	5	2011	HDD	Chicago	11	1	2							2
2010	HDD	NY	12	1	2	-	-	-	-	2	2011	HDD	Chicago	12	1	2							2
2011	HDD	NY	1	1	4	1	-	-	-	5	2004	CDD	Dallas	9	1	1							1
2011	HDD	NY	2	1	5	-	-	-	-	5	2005	CDD	Dallas	8	1	1							1
2011	HDD	NY	3	1	1	1	-	-	-	2	2006	CDD	Dallas	6	1	1							1
2011	HDD	NY	11	1	1	-	-	-	-	1	2006	CDD	Dallas	9	1	1							1
2011	HDD	NY	12	1	2	-	-	-	-	2	2007	CDD	Dallas	7	1	1							1
2012	HDD	NY	2	1	1	-	-	-	-	1	2008	CDD	Dallas	5	1	6							6
2004	CDD	NY	9	1	-	1	-	-	-	1	2008	CDD	Dallas	6	1	1							1
2005	CDD	NY	5	1		1	-	-	-	1	2008	CDD	Dallas	7	1	2							2
2005	CDD	NY	6	1		8	1	-	-	9	2008	CDD	Dallas	8	1	1							1
2005	CDD	NY	7	1			-	-	-	1	2009	CDD	Dallas	5	1	4	1						5
2005	CDD	NY	8	1		4	-	-	-	4	2009	CDD	Dallas	6	1	1		1					2
2006	CDD	NY	6	1		4	-	-	-	4	2009	CDD	Dallas	8	1	1							1
2006	CDD	NY	7	1		1	-	-	-	1	2009	CDD	Dallas	9	1	1							1
2006	CDD	NY	8	1		3	-	-	-	3	2010	CDD	Dallas	5	1	1							1
2006	CDD	NY	9	1		3	-	-	-	2	2010	CDD	Dallas	6	1	2	3						5
2007	CDD	NY	7	1		1	-	-	-	1	2010	CDD	Dallas	7	1		2						2
2007	CDD	NY	8	1		1	-	-	-	1	2010	CDD	Dallas	8	1		2						2
2007	CDD	NY	9	1		1	-	-	-	1	2010	CDD	Dallas	9	1		3	1					4
2008	CDD	NY	6	1		5	-	-	-	5	2011	CDD	Dallas	5	1			1					1
2008	CDD	NY	7	1		1	-	-	-	1	2011	CDD	Dallas	6	1	3	1						4
2008	CDD	NY	8	1		1	-	-	-	1	2011	CDD	Dallas	7	1	3							3
2009	CDD	NY	8	1		2	-	-	-	2	2011	CDD	Dallas	8	1	2							2
2010	CDD	NY	5	1		2	-	-	-	2	2012	CDD	Dallas	2	1	2							2
2010	CDD	NY	5	1		2	-	-	-	2													
2010	CDD	NY	6	1		1	-	-	-	1													
2010	CDD	NY	7	1		1	-	-	-	1													
2010	CDD	NY	9	1		1	-	-	-	1													
2011	CDD	NY	9	1		1	-	-	-	1													
2011	CDD	NY	6	1		3	-	-	-	3													

Figure 11: New York monthly HDD option prices with time to maturity in one month against strike prices and trading days in March 2006, HDD Atlanta Seasonal option prices with time to maturity in 6 months and trading dates in October 2007, Dallas monthly CDD options with time to maturity in one month traded in June 2010.



scale:

$$R^2 = 1 - \frac{\sum_{i=1}^{N_i} \sum_{j=1}^{N_i} \sum_{k=1}^{N_{ij}} \{\log y_{ijk} - \log C_{ij}^N(w, \theta)\}^2}{\sum_{i=1}^{N_i} \sum_{j=1}^{N_i} \sum_{k=1}^{N_{ij}} \log y_{ijk}^2}$$

When R^2 is close to one, model prices are close to market prices and the model produces nice model fit. In the Bayesian quadrature method, we calculate R^2 at each swipe of the Markov chain Monte Carlo algorithm, and summarize its posterior mean and quantiles for inference. Because all the mean, median, and the 2.5% and 97.5% quantiles of the R^2 calculated in the MCMC algorithm are close to one, we conclude that the Bayesian quadrature method produces an almost perfect fit for all these trading days. Figure 12 presents dynamics of the implied SPD, all of them deviating from log-normality. However, skewness and kurtosis of weather options can be either positively or negatively skewed depending on futures maturity. This is different with what documented in index options market, Bakshi et al. (2010), interest rate derivatives market Li and Zhao (2009) and in temperature markets Benth et al. (2007). SPDs are depending on the conditional volatility, the SPD is wider when the conditional volatility is high than when it is low. This is also explained by the economic behavior of agents sensitive to weather conditions. Investors expect that temperature variations, that affect their cash flows, will occur with high probability in winter times than in summer times (except for WD for Australia). WD-SPD's tend to be positively skewed for short maturity contracts and slightly negatively skewed for long maturity contracts as the option prices only provides information for the right tail of the SPD. The evolution of the parameters of the quadrature method is displayed in Figure 13.

A simple way to investigate the dynamics of the implied SPD at each trading day is to calculate moments based on the quadrature method. In each swipe of the Markov chain Monte Carlo algorithm, we calculate the mean (μ), volatility (ν), skewness (s), and kurtosis (κ) of the quadrature method, by the following formulas,

$$\begin{aligned}\mu &= \sum_{n=1}^N w_n \theta_n \\ \nu &= \sqrt{\sum_{n=1}^N w_n (\theta_n - \mu)^2} \\ s &= \sum_{n=1}^N w_n (\theta_n - \mu)^3 / \nu^3 \\ \kappa &= \sum_{n=1}^N w_n (\theta_n - \mu)^4 / \nu^4\end{aligned}$$

Table 4 shows the posterior means of these four quantities of the quadrature method.

4 Conclusions

We estimate SPDs for WDs using the Bayesian quadrature method. The WD market is characterized by its incompleteness and illiquid options. This makes the estimation of

Figure 12: Quadrature method and smoothed SPDs implied from New York monthly HDD option prices with time to maturity in one month against strike prices and trading days in March 2006, HDD Atlanta Seasonal option prices with time to maturity in 6 months and trading dates in October 2007, Dallas monthly CDD options with time to maturity in one month traded in June 2010.

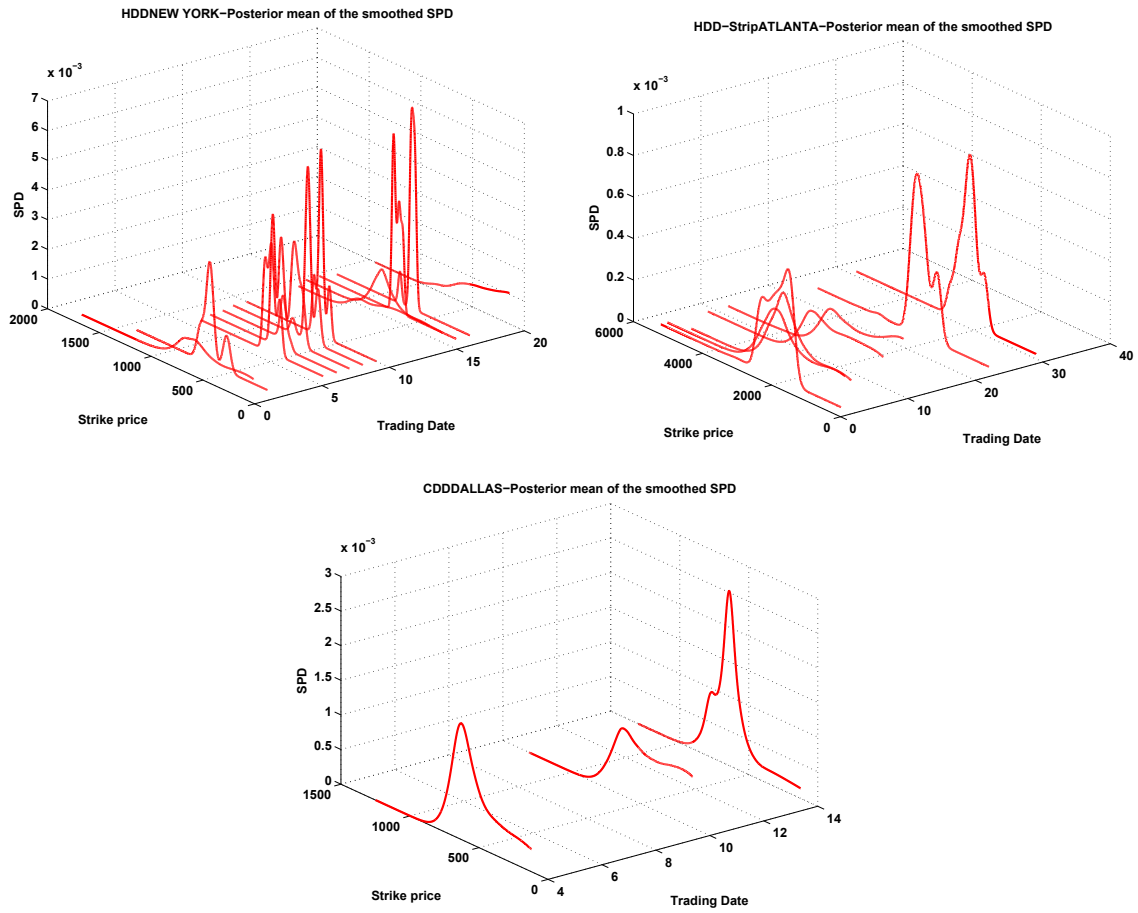


Figure 13: Dynamics of the parameter of the quadrature method implied from New York monthly HDD option prices with time to maturity in one month against strike prices and trading days in March 2006, HDD Atlanta Seasonal option prices with time to maturity in 6 months and trading dates in October 2007, Dallas monthly CDD options with time to maturity in one month traded in June 2010.

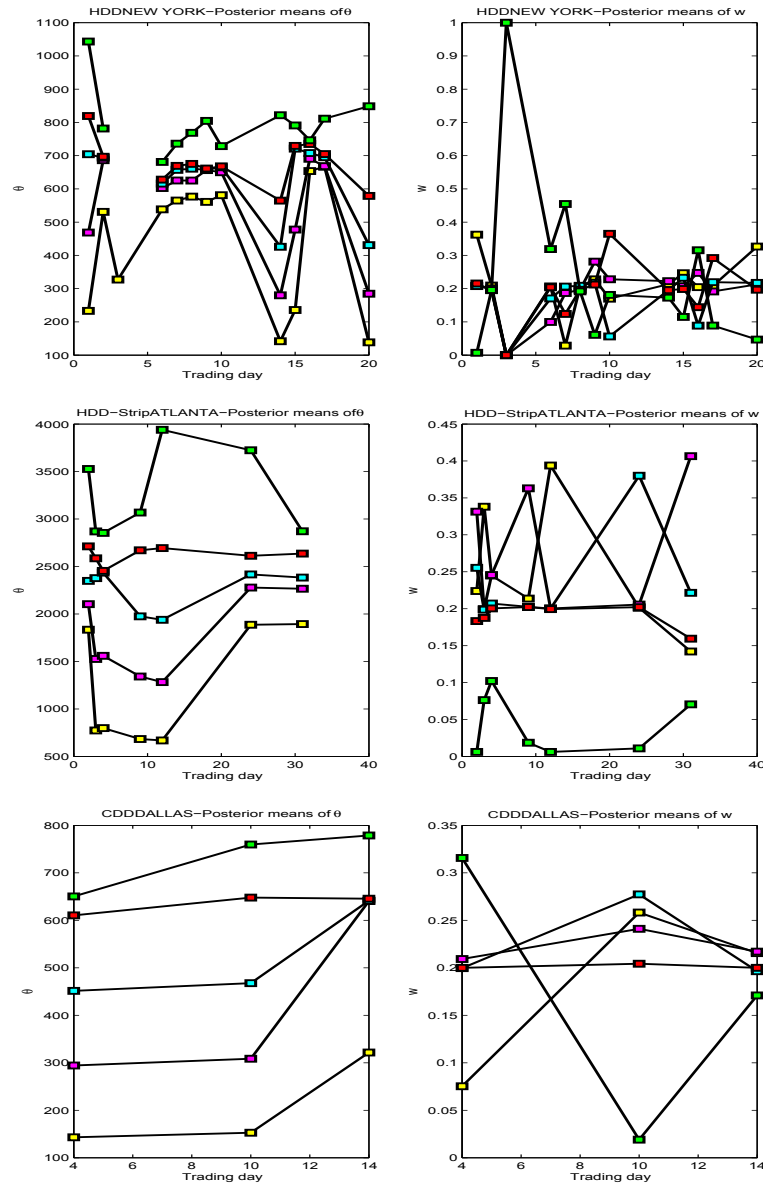


Table 4: Posterior mean of the mean (μ), volatility (v), skewness (s), and kurtosis (κ) of the quadrature method at each trading day (TD) calibrated from New York HDD monthly options in March 2006, Atlanta HDD seasonal strip options in October 2007 and Dallas CDD monthly options in June 2010.

	TD	2	3	6	7	8	9	10	14	15	16	17	20
HDD	μ	511.16	676.00	618.80	676.31	660.00	645.00	660.00	422.78	554.10	708.58	698.31	380.26
NY	v	248.98	81.57	57.02	73.88	61.38	59.89	45.07	240.72	217.04	41.54	42.11	205.93
	s	0.09	-0.92	0.25	4.30	0.48	0.21	0.38	0.43	-0.42	1.47	1.65	0.60
	κ	1.54	2.70	2.79	56.28	2.26	3.28	2.98	2.16	1.70	5.66	7.59	2.90
	TD	2	3	4	9	12	24	31					
HDD	μ	2225.54	1741.96	1865.99	1630.19	1470.13	2335.12	2340.70					
Strips	v	324.45	818.35	757.39	751.52	835.64	287.44	258.52					
Atlanta	s	0.58	-0.05	-0.33	0.21	0.59	0.64	0.07					
	κ	3.44	1.42	1.73	2.15	2.44	7.59	3.00					
	TD	4	10	14									
CDD	μ	489.96	390.42	596.63									
Dallas	v	175.79	195.06	155.40									
	s	-0.83	0.18	-0.57									
	κ	2.72	1.98	2.76									

the SPD a statistical challenge. However, the quadrature method, in advantage to the parametric and other non-semiparametric techniques, avoids model miss-specification and allows the SPD estimation by a parsimonious model. The techniques is computationally fast and robust. We present empirical results on real CME temperature derivative data, which help us to understand the dynamics of SPD. The results suggest that the SPD of weather derivatives exhibits a non-normal behavior type.

References

- Abadir, K. and Rockinger, M. (2003). Density functionals, with an option-pricing application. *Econometric Theory*, 19(5):778–811.
- Ait-Sahalia, Y. and Duarte, J. (2003). Nonparametric option pricing under shape restrictions. *Journal of Econometrics*, 116:9–47.
- Ait-Sahalia, Y. and Lo, A. W. (1998). Nonparametric estimation of state-price densities implicit in financial asset prices. *Journal of Finance*, 53:499–547.
- Ait-Sahalia, Y. and Lo, A. W. (2000). Nonparametric risk management and implied risk aversion. *Journal of Econometrics*, 94:9–51.
- Bakshi, G., Madan, D., and g. Panayotov (2010). Returns of claims on the upside and the viability of u-shaped pricing kernels. *Journal of Financial Economics*, 97(1):130–154.
- Benth, F., Benth, S., and Koekebakker, S. (2007). Putting a price on temperature. *Scandinavian Journal of Statistics*, 34:746–767.
- Benth, F., Härdle, W. K., and López-Cabrera, B. (2011). *Pricing Asian temperature risk in Statistical Tools for Finance and Insurance. 2nd. edition (Cizek, Härdle and Weron, eds.)*. Springer Verlag Heidelberg.

- Breeden, D. and Litzenberger, R. (1978). Price for state-contingent claims implicit in option prices. *Journal of Business*, 51:621–651.
- Casella, G. and Berger, R. L. (2001). *Statistical Inference*. Duxbury Press.
- Chen, M. and Shao, Q. (1997). On Monte Carlo methods for estimating ratios of normalizing constants. *The Annals of Statistics*, 25 (4):1563–1594.
- Derman, E. and Kani, I. (1994). Riding on the smile. *Risk*, 7:32–39.
- Dupire, B. (1994). Pricing with a smile. *Risk*, 7:18–20.
- Fan, J. and Mancini, L. (2009). Option pricing with model-guided nonparametric methods. *Journal of the American Statistical Association*, 104 (488):1351–1372.
- Garcia, R., Ghysels, E., and Renault, E. (2010). *Econometrics of Option Pricing Models in Handbook of Financial Econometrics (Y. Ait-Sahalia and L.P. Hansen, eds.)*, volume 1. North Holland.
- Ghysels, E., Harvey, A., and Renault, R. (1995). *Stochastic Volatility in Handbook of Statistics 14, Statistical Methods in Finance (G.S. Maddala and C.R. Rao, eds.)*. North Holland, Amsterdam.
- Ghysels, E., Patilea, V., Renault, E., and Torres, O. (1997). *Nonparametric methods and option pricing in Statistics in Finance (D. Hand and S. Jacka, eds.)*. Edward Arnold, London.
- Giacomini, R., Gottschling, A., Haefke, C., and White, H. (2008). Mixtures of t distributions for finance and forecasting. *Journal of Econometrics*, 144:175–192.
- Härdle, W., Müller, M., Sperlich, S., and Werwatz, A. (2004). *Nonparametric and Semiparametric Models*. Springer Verlag.
- Härdle, W. K. and Hlávka, Z. (2009). Dynamics of state price densities. *Journal of Econometrics*, 150:1–15.
- Härdle, W. K. and López-Cabrera, B. (2012). Inferring the market price of weather risk. *Applied Mathematical Finance*, 19 (1):59–95.
- Jackwerth, J. and Rubinstein, M. (1996). Recovering probabilities distributions from option prices. *Journal of Finance*, 51:1611–1631.
- Li, H. and Zhao, F. (2009). Nonparametric estimation of state-price densities implicit in interest rate cap prices. *Review of Financial Studies*, 22(11):4335–4376.
- Renault, E. (1997). *Econometric models of options pricing errors in Advances in Economics and Econometrics: Theory and Applications * Seventh World Congress, Vol. III, (D. Kreps, K. Wallis, eds.)*. Cambridge University Press, Cambridge.
- Rosenberg, J. and Engle, R. (2002). Empirical pricing kernels. *Journal of Financial Econometrics*, 64:1611–1631.

- Rubinstein, M. (1994). Implied binomial trees. *Journal of Finance*, 49:771–818.
- Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*. Chapman and Hall.
- Teng, H.-W. and Liechty, J. (2009). Bayesian quadrature approaches to state price density estimation. *working paper*. Available at <http://www.personal.psu.edu/jcl12/Bayesian%20Quadrature.pdf>.
- Ueberhuber, C. W. (1997). *Numerical Computation 2: Methods, Software, and Analysis*. Springer-Verlag.
- Yatchew, A. and Härdle, W. (2006). Nonparametric state price density estimation using constrained least squares and the bootstrap. *Journal of Econometrics*, 133(2):579–599.
- Yuan, M. (2009). State price density estimation via nonparametric mixtures. *Annals of Applied Statistics*, 3(3):963–984.

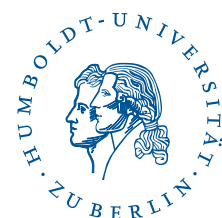
SFB 649 Discussion Paper Series 2013

For a complete list of Discussion Papers published by the SFB 649, please visit <http://sfb649.wiwi.hu-berlin.de>.

- 001 "Functional Data Analysis of Generalized Quantile Regressions" by Mengmeng Guo, Lhan Zhou, Jianhua Z. Huang and Wolfgang Karl Härdle, January 2013.
- 002 "Statistical properties and stability of ratings in a subset of US firms" by Alexander B. Matthies, January 2013.
- 003 "Empirical Research on Corporate Credit-Ratings: A Literature Review" by Alexander B. Matthies, January 2013.
- 004 "Preference for Randomization: Empirical and Experimental Evidence" by Nadja Dwenger, Dorothea Kübler and Georg Weizsäcker, January 2013.
- 005 "Pricing Rainfall Derivatives at the CME" by Brenda López Cabrera, Martin Odening and Matthias Ritter, January 2013.
- 006 "Inference for Multi-Dimensional High-Frequency Data: Equivalence of Methods, Central Limit Theorems, and an Application to Conditional Independence Testing" by Markus Bibinger and Per A. Mykland, January 2013.
- 007 "Crossing Network versus Dealer Market: Unique Equilibrium in the Allocation of Order Flow" by Jutta Dönges, Frank Heinemann and Tijmen R. Daniëls, January 2013.
- 008 "Forecasting systemic impact in financial networks" by Nikolaus Hautsch, Julia Schaumburg and Melanie Schienle, January 2013.
- 009 "'I'll do it by myself as I knew it all along': On the failure of hindsight-biased principals to delegate optimally" by David Danz, Frank Hüber, Dorothea Kübler, Lydia Mechtenberg and Julia Schmid, January 2013.
- 010 "Composite Quantile Regression for the Single-Index Model" by Yan Fan, Wolfgang Karl Härdle, Weining Wang and Lixing Zhu, February 2013.
- 011 "The Real Consequences of Financial Stress" by Stefan Mittnik and Willi Semmler, February 2013.
- 012 "Are There Bubbles in the Sterling-dollar Exchange Rate? New Evidence from Sequential ADF Tests" by Timo Bettendorf and Wenjuan Chen, February 2013.
- 013 "A Transfer Mechanism for a Monetary Union" by Philipp Engler and Simon Voigts, March 2013.
- 014 "Do High-Frequency Data Improve High-Dimensional Portfolio Allocations?" by Nikolaus Hautsch, Lada M. Kyj and Peter Malec, March 2013.
- 015 "Cyclical Variation in Labor Hours and Productivity Using the ATUS" by Michael C. Burda, Daniel S. Hamermesh and Jay Stewart, March 2013.
- 016 "Quantitative forward guidance and the predictability of monetary policy – A wavelet based jump detection approach –" by Lars Winkelmann, April 2013.
- 017 "Estimating the Quadratic Covariation Matrix from Noisy Observations: Local Method of Moments and Efficiency" by Markus Bibinger, Nikolaus Hautsch, Peter Malec and Markus Reiss, April 2013.
- 018 "Fair re-valuation of wine as an investment" by Fabian Y.R.P. Bocart and Christian M. Hafner, April 2013.
- 019 "The European Debt Crisis: How did we get into this mess? How can we get out of it?" by Michael C. Burda, April 2013.

SFB 649, Spandauer Straße 1, D-10178 Berlin
<http://sfb649.wiwi.hu-berlin.de>

This research was supported by the Deutsche
Forschungsgemeinschaft through the SFB 649 "Economic Risk".



SFB 649 Discussion Paper Series 2013

For a complete list of Discussion Papers published by the SFB 649, please visit <http://sfb649.wiwi.hu-berlin.de>.

- 020 "Disaster Risk in a New Keynesian Model" by Maren Brede, April 2013.
- 021 "Econometrics of co-jumps in high-frequency data with noise" by Markus Bibinger and Lars Winkelmann, May 2013.
- 022 "Decomposing Risk in Dynamic Stochastic General Equilibrium" by Hong Lan and Alexander Meyer-Gohde, May 2013.
- 023 "Reference Dependent Preferences and the EPK Puzzle" by Maria Grith, Wolfgang Karl Härdle and Volker Krätschmer, May 2013.
- 024 "Pruning in Perturbation DSGE Models - Guidance from Nonlinear Moving Average Approximations" by Hong Lan and Alexander Meyer-Gohde, May 2013.
- 025 "The 'Celtic Crisis': Guarantees, transparency, and systemic liquidity risk" by Philipp König, Kartik Anand and Frank Heinemann, May 2013.
- 026 "State Price Densities implied from weather derivatives" by Wolfgang Karl Härdle, Brenda López-Cabrera and Huei-Wen Teng, May 2013.

SFB 649, Spandauer Straße 1, D-10178 Berlin
<http://sfb649.wiwi.hu-berlin.de>

This research was supported by the Deutsche
Forschungsgemeinschaft through the SFB 649 "Economic Risk".

